PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/62, C07K 19/00, C12P 21/02, C12N 15/70, 15/74, 15/81, 15/80, 1/21, 1/15, 1/16

A1

(11) International Publication Number: **WO 98/56928**

(43) International Publication Date: 17 December 1998 (17.12.98)

(21) International Application Number: PCT/NL98/00335

(22) International Filing Date: 8 June 1998 (08.06.98)

(30) Priority Data:
97201790.9 13 June 1997 (13.06.97) EP
(34) Countries for which the regional or international application was filed: NL et al.

(71) Applicant *(for all designated States except US)*: STICHTING INSTITUUT VOOR DIERHOUDERIJ EN DIERGEZONDHEID [NL/NL]; Edelhertweg 15, NL–8219 PH Lelystad (NL).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: HARMSEN, Michiel, Marie [NL/NL]; Terletstraat 91, NL–1107 RL Amsterdam (NL). DE GEUS, Bernard [NL/NL]; Mondlanestraat 25, NL–2807 RJ Gouda (NL).

(74) Agent: SMULDERS, Th., A., H., J.; Vereenigde Octrooibureaux, Nieuwe Parklaan 97, NL–2587 BN The Hague (NL).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: INCREASED PRODUCTION OF PROTEINS BY USING CHAPERONE–LIKE PROTEINS

(57) Abstract

The invention relates to the production of recombinant proteins that are expressed in heterologous host cells and to methods to increase production levels of these proteins by the use of a chaperon–like protein. Producing or isolating a protein or polypeptide in large amounts and in a relatively pure form from its original (homologous) host or host cell where it is expressed is cumbersome and often economically unattractive. However, the development of recombinant DNA methodology has opened the possibility to produce proteins by expression, e.g. in a heterologous host or host cell. A wide and still expanding array of foreign polypeptides or proteins can now be produced or expressed in heterologous expression systems. The invention provides a method to express a polypeptide encoded by a nucleic acid sequence present in a host cell which host cell additionally comprises a chaperone–like protein capable of specific binding with said polypeptide.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

Title: Increased production of proteins by using chaperone-like proteins.

The invention relates to the production of recombinant proteins that are expressed in heterologous host cells and to methods to increase production levels of these proteins.

Producing or isolating a protein or polypeptide in large amounts and in a relatively pure form from its original (homologous) host or host cell where it is expressed is cumbersome and often economically unattractive. However, the development of recombinant DNA methodology has opened the possibility to produce proteins by (over)expression, e.g. in a heterologous host or host cell. A wide and still expanding array of foreign polypeptides or proteins can now be produced or expressed in homologous or heterologous expression systems; as examples can be listed erytropoetin, chymosin, transferrin, and many more are known.

Expression systems, which can be *in vitro* expression systems or *in vivo* expression systems comprising hosts cells, are frequently used for the production or expression of a foreign polypeptide or protein originate from at least five kingdoms: such as (1) prokaryotes, such as *Escherichia coli* (2) fungi, including the yeasts, such as *Saccharomyces cerevisiae,* and the filamentous fungi, such as *Aspergillus awamori,* (3) mammals, including intact organisms such as *Bos tauris,* and various mammalian cell lines originating from several species, (4) insect cell lines, such as *Spodoptera frugiperda,* and (5) plants (Hodgson, 1993, Bio/Technology 11: 887–893). Hosts from the first two kingdoms can most easily be genetically modified and have a potential for high-level polypeptide or protein production at low cost.

However, the success with which foreign proteins can be produced in a heterologous organism is highly variable. Simple proteins can often be expressed reasonably well in host cells derived from for instance the prokaryotes or the yeasts and fungi, however, more complicated proteins are often refractory to expression in these simple host cells, and often require expression in the more complicated insect, plant or animal cell systems, or even require expression in very specialized cells or even in whole organisms which comprise the specialized cells and cell systems necessary for proper expression and processing of the wanted protein. In this light it must be noted that several of the early promises of biotechnology relating to the production of complicated proteins or polypeptides in relatively simple expression systems at low costs have not yet materialized.

Heterologous (over)expression comprises several steps which all have a distinct impact on the success with which a foreign protein can be produced.

First, the gene encoding the protein of interest or wanted protein must be isolated from the original host.

Second, this gene must be redesigned and inserted into a vector suitable for expression in the heterologous host. This often includes replacement of the transcription and translation signals of the pertinent gene by sequences which originate from the heterologous host in order to obtain high levels of mRNA which can be translated. This may also include removal of introns and change of codon usage. With current state of the art of recombinant DNA technology it is possible to perform these first two steps for any protein (e.g. see the reviews of Romanos et al., 1992, Yeast 8: 423-488; Wall and Pluckthun, 1995, Curr. Opin. Biotechnol. 6: 507-516).

Third, the heterologous protein must contain the appropriate targeting signals for directing the protein to the desired subcellular compartment. Especially targeting to

the secretory pathway is frequently used since the purification of the foreign protein once it is secreted into the culture medium is relatively simple. For this purpose it may be necessary to remove the original targeting signals and introduce new targeting signals. With current knowledge of targeting signals this can also be performed for any protein.

Fourth, the protein must be post-translationally modified in a correct manner. Especially *N*-glycosylation of proteins is an important modification in this respect since many commercially interesting proteins need to be *N*-glycosylated correctly to attain biological activity. Since *N*-glycosylation of proteins does not occur in prokaryotes and occurs in a different manner in the cells or organisms originating from the various kingdoms, the production of mammalian proteins, which form the majority of commercially interesting proteins, in many of these organisms often results in absent or improper *N*-glycosylation. Therefore, the production of higher eukaryotic proteins in prokaryotes or yeasts and fungi is limited to proteins that do not require *N*-glycosylation (Parekh, 1994, Biologicals 22: 113-119). Furthermore, proteins that are normally not glycosylated because they originate from compartments where this modification normally does not occur, such as cytosolic proteins, may become fortuitously glycosylated upon production in the eukaryotic secretory pathway, which may result in inhibition of their biological function. However, part of the problems related to glycosylation can be resolved by careful selection of an appropriate expression system such as insect cells or plant cells, by choosing amino acid sequences that avoid glycosylation sites, or by further treatment of the expressed protein with glycosidase that help trim the carbodydrate side chains of the protein to normal proportions.

Finally, the polypeptide must fold into the correct, authentic conformation, which can be secreted and is biologically active. This bottle-neck is the step which hampers the commercial production of foreign proteins most (see Cleland, 1993, Impact of protein folding on biotechnology, in: Cleland (Ed.), Protein folding: in vivo and in vitro, American Chemical Society, Washington, pp. 1-21). Especially for frequently used heterologous hosts such as *E. coli* and *S. cerevisiae* it is well-documented that overexpression of heterologous proteins results in the formation of folding-intermediates of these proteins and insoluble aggregates of these proteins which as a result are biologically inactive (Edgington, 1992, Bio/Technology 10, 1413-1420; Tuite and Freedman, 1994, Trends Biotechnol. 12, 432-435). In other words, although the expression system can easily translate the foreign nucleic acid sequence and express the protein in large quantities, the further routing of the protein is greatly hampered by mis-folding and thus the formation and intra-cellular accumulation of insoluble aggregates, finally constipating the protein machinery of the host cell.

The primary sequence of a protein determines the nature of the fold that the protein will assume, as is indicated by the fact that many proteins can regain their native conformation after spontaneous *in vitro* refolding (Anfinsen, 1973, Science 181: 223-230). However, the efficiency of this process *in vitro* is often low, because the transient exposure of hydrophobic surfaces of immature forms or folding intermediates during the process of refolding results in aggregation, which is an irreversible process that competes with the formation of the correct fold.

Protein folding *in vivo* (in the cell) is now known to be assisted by a diverse set of naturally occurring (native) proteins collectively known as foldases (e.g. see reviews by

Gething and Sambrook, 1992, Nature 355: 33-45; Hartl, 1996, Nature 381: 571-580). A major subgroup of the foldases are the molecular chaperones, which are defined as a group of unrelated classes of proteins that bind to and stabilize an otherwise unstable conformer or immature form of another protein and by controlled binding and release facilitate its correct fate *in vivo*, be it folding, oligomeric assembly, transport to a particular subcellular compartment, or disposal by degradation. Cells have different compartments with different physiological conditions resulting in different requirements for protein folding and, therefore, different chaperone contents. Gram negative prokaryotes such as *E. coli* contain two compartments which are important in this respect: the cytosol and the periplasmic space. Eukaryotes contain several compartments where protein folding occurs, of which the most important ones are (1) the cytosol, (2) the endoplasmic reticulum (ER), (3) mitochondria and (4) chloroplasts. The cytosol of both prokaryotes and eukaryotes is a relatively reducing environment as compared to the ER.

Expression of proteins in a compartment from which they do not originate often results in misfolding, either due to an inappropriate oxidation state or due to the lack of the specific foldases or chaperones required for folding. Even when proteins are expressed in a homologous compartment of a heterologous host they often misfold. Chaperones *per se* prevent incorrect interactions between polypeptides, thus typically increasing the yield but not the rate of folding reactions. This distinguishes them from the so-called folding catalysts, protein disulphide isomerases and peptidyl-prolyl isomerases. These enzymes catalyse the formation and rearrangement of disulphide bonds in secretory proteins and the *cis-trans* isomerization of peptide bonds preceding proline residues, respectively. Chaperones can bind to an immature form of a specific protein and shield the protein

from the cellular environment, thereby inducing or catalysing proper folding and conformation of to-be-expressed proteins, and preventing mis-folding and aggregation. In nature, proteins with such activity can be found in many protein families which have different functions, such as the Hsp60, Hsp70, Hsp90, Bip, chaperonin and Calnexin. All these foldases co-operate in a complex interplay that is regulated by other proteins during protein folding.

Foreign proteins can occasionally be produced in a mis-folded state in a heterologous host and subsequently isolated after e.g. disruption of the host cells. Also, occasionally, a mis-folded protein is secreted from the host cell despite its misfolded state. Protocols have been designed to refold the misfolded proteins *in vitro*, however, this is a cumbersome process that is costly and does not result in a high yield of the wanted protein. Several of the *in vitro* folding protocols entail the use of antibodies as the *in vitro* folding agent. However, selection of these antibodies for proper folding activity is difficult and only in a few occasions possible, mainly because the reactivity of the antibodies is directed against the proper folded protein and against the unknown conformation of the mis-fold protein and not directed against the immature form from which the correct form should follow.

Attempts to improve *in vivo* foreign protein production and secretion from the host cell by co-expression of a corresponding chaperone protein (thereby using foldases that have by nature been selected to properly function as chaperone and having a substrate specificity corresponding to the foreign protein) originating from the homologous host have been unsuccessful (Hsu et al., 1994, Protein Expression Purification 5: 595-603) or resulted in only minor improvements, partly due to the inability of the heterologous host to produce the chaperone itself in a biologically active

form. Furthermore, heterologous proteins are usually expressed at a high level, which results in overloading of an already in itself not well-suited chaperone machinery.

There are several examples of attempts to stimulate *in vivo* foreign protein production and secretion from the host cell by over-expression of a chaperone which is already normally expressed in the heterologous host or expression system. Several reports (Goloubinoff et al., 1989, Nature 337: 44-47; Phillips and Silhavy, 1990, Nature 344, 882-884; Lee and Olins, 1992, J. Biol. Chem 267: 2849-2852; Harmsen et al., 1996: Appl. Microbiol. Biotechn. 46, 365-370) are indeed showing that a limited protein folding capacity of the heterologous host is responsible for the inefficient production. However, attempts to remedy this limited folding capacity by over-expressing homologous foldases generally fails (Knappik et al., 1993, Bio/Technology 11, 77—83; Harmsen et al., 1996, Appl. Microbiol. Biotechnol. 46: 365-370; Van Gemeren, 1997, PhD thesis, University of Utrecht), obviously due to the lack of substrate specificity of the tentative chaperone and may even result in a decrease in total production (Lah et al., 1994, Hum. Antibod. Hybridomas 5: 48-56), dependent on the particular heterologous protein studied.

In overview, reasons why over-expression of foldases generally do not stimulate foreign protein folding and secretion from the host cell are several fold. Firstly, the foldase itself may misfold upon over-expression, either due to its heterologous nature whereby the chaperone-like properties of the foldase do not correspond to the requirements of the host cell, or due to its over-expression. Secondly, it is more than likely that the particular foldase chosen for over-expression does not function for the heterologous protein. Thirdly, since protein folding is assisted by a complex machinery consisting of many foldases

it is impossible to stimulate the process by simply over-expressing a single and ill-fitting component.

The present invention provides a method of protein production in a expression system by providing a chaperone-like protein also referred to as design chaperone. The invention further provides a chaperone-like protein with at least one binding site with binding activity for at least one immature form of a polypeptide for which it originally had no, less or different binding activity. A chaperone-like protein is a non-native chaperone or other binding protein specifically devised and brought into being to be capable of the (transient) binding with a protein that is done in nature by a chaperone. A chaperone-like protein is selected to have a distinct substrate specificity and other chaperone-like properties and is deliberately generated and tailored via recombinant techniques to have the capacity to cause proper folding, transporting or secretion of a protein that e.g. is being expressed in an expression system. Binding of a chaperone-like protein to the wanted polypeptide takes place by the more-or-less transient formation of multiple non-covalent bonds between the two proteins. Although the attractive forces involved in these bonds are weak by comparison with covalent bonds, the multiplicity of the bonds leads to a often considerable binding energy. These forces are likely to play a vital role in the folding and re-folding processes that are needed to arrive, via various folding-intermediates or immature forms, at a correctly folded polypeptide that can be expressed and secreted, since any variation in form or shape will cause a variation in the total binding energy between the two proteins. As binding proteins, a wide array of proteins is known in the art, such as receptor molecules, chaperone proteins, polyclonal or monoclonal (synthetic) antibodies, minibodies, binding peptides, 'phage' antibodies derived via phage display

techniques, single-chain antibody-variable fragments (scVF's), heavy-chain antibody fragments, in general all proteins belonging to the immunoglobuline superfamily (including bacterial proteinases, MHC molecules, T-cell receptors), and so on. A chaperone-like protein generally has a narrow band of specificity, in that it is specifically reactive with the polypeptide for which it is designed and with structurally, functionally or antigenically related proteins. The invention provides a method to express a wanted polypeptide encoded by a nucleic acid sequence or molecule present in a host cell which additionally comprises a chaperone-like protein capable of specific binding with said wanted polypeptide. Said chaperone-like protein used in the method provided by the invention comprises at least one binding site reactive with the wanted polypeptide. Such a binding site of a chaperone-like protein provided by the invention can be derived from a wide array of sources. The invention for example provides a chaperone-like protein comprising a binding site that is derived from a (synthetic) peptide sequence that is reactive with the wanted polypeptide. Such a peptide can be selected by various methods known in the art, such as PEPSCAN methods, applied molecular evolution methods, PEPSCAN replacement mapping, and so on. The invention for example also provides a chaperone-like protein comprising a binding site that is derived from naturally occurring binding proteins such as enzymes, receptor molecules, cell surface proteins, chaperones, transfer proteins, and so on. A preferred embodiment of the invention entails the use of a chaperone-like protein that comprises a binding site that is derived from an antibody, an antibody fragment, a single-chain antibody fragment, or from any the variable regions of the heavy and/or light chain of antibodies. In the experimental part of the description, binding sites derived from heavy-chain antibody fragments of

camelids are used, but it is within the ordinary skills of those working in the field of immunology to select an antibody derived from an appropriate species, varying from poultry to sharks, providing (a) binding site(s) for use in a chaperone-like protein according to the invention. A chaperone-like protein provided by the invention is instrumental in causing improved folding of the protein for which it bears specific binding activity. The invention also provides an expression system provides with a chaperone-like protein. Such an expression system can be an *in vivo* system, using host cells of various origin, or an *in vitro* system, using e.g. lysates of various host cells, such as reticulocyte lysates or wheat germ extracts in which *in vitro* (transcription-)translation can occur. In such an expression system provided by the invention the chaperone-like protein enhances correct folding of the expressed polypeptide. Furthermore, when present in a host cell expressing a polypeptide for which it bears specific binding activity, a chaperone-like protein is instrumental in causing improved routing through the cell and an increase of secretion of the properly folded polypeptide from said host cell. The invention also provides a nucleic acid molecule or molecules encoding a chaperone-like protein. Such a nucleic acid molecule provided by the invention can be both DNA or RNA, be it single or double stranded, and can be subjected to recombinant techniques known in the art. Such nucleic acid molecules can additionally comprise a nucleic acid sequence encoding the wanted polypeptide. The invention also provides vectors, expression systems and host cells comprising such a nucleic acid. Such a nucleic acid molecule can for example be place into an expression vector. Such a vector may be replicable or non-replicable, and can be integrated partly or whole in the chromosomal DNA of the host cell or the expression system. In a particular embodiment of the

invention, such a vector may comprise the sequences encoding both the wanted polypeptide and the chaperone-like protein. A vector can easily be swapped from host cell to e.g. an *in vitro* expression system, or to yet another host cell, and vice versa. A preferred embodiment of the invention is a vector containing sequences allowing for controlled expression of the chaperone-like protein and/or the wanted polypeptide. The invention provides host cells comprising a nucleic acid encoding a chaperone-like protein which is preferably linked to a promotor, enhancer, upstream control elements, transcription factors, repressor binding sites, polyadenylation sites, initiation site and so on, thereby e.g. facilitating constitutive or, alternatively, inducible expression of the chaperone-like protein. The invention thus allows various modes of expression of the chaperone-like protein, in relation to the time and mode of expression of the wanted polypeptide. The invention for example provides a method whereby the chaperone-like protein is being expressed in advance to, or at the same time as, the expression of the wanted polypeptide, thereby making functional chaperone-like protein available in the host cell at the times most needed, namely when the wanted polypeptide is being produced. The invention also provides a method of protein production whereby co-expressing a chaperone-like protein causes improved folding and/or secretion of a protein that is expressed in an expression system. The invention provides a method whereby the chaperone-like protein is being co-expressed together with its protein, which is either autogenous or heterogenous to the expression system. A preferred embodiment of the invention provides a method to produce a protein in a heterologous expression system whereby a chaperone-like protein is being co-expressed, said chaperone-like protein being capable of facilitating proper folding and/or secretion of the protein. Also, the method

provides host cells expressing a foreign protein and equipped with a nucleic acid molecule encoding a chaperone-like protein according to the invention. In the host cells provided by the invention the folding of the wanted foreign protein is helped by the specific chaperone-like protein, preventing the formation of accumulating aggregates of misfolded protein and thereby easing the routing and secretion of the foreign protein out of the cell. Host cells provided by the invention comprise a nucleic acid molecule or molecules encoding a chaperone-like protein, which in one embodiment of the invention can be expressed transiently, in another embodiment constitutively. In yet another embodiment of the invention the expression of the chaperone-like protein is inducible. Furthermore, host cells provided by yet another embodiment of the invention comprise multiple copies of a nucleic acid molecule or molecules encoding one or more chaperone-like proteins. Also, said nucleic acid molecule encoding a chaperone-like protein can be integrated in the genome of the host cell, or alternatively, can be part of a vector (such as a plasmid, or shuttle vector or bacteriophage or cosmid or virus) with which the host cell has been infected or transformed. The invention provides host cells comprising a chaperone-like protein. Such host cells can be prokaryotic, for example bacteria belonging to the genera *Escherichia, Bacillus, Streptomyces and Lactobacillus*. The invention also provides other host cells comprising a chaperone-like protein, e.g, wherein the host cell is a fungus, for example selected from the group of yeasts belonging to the genera *Saccharomyces, Kluyveromyces, Hansenula, Pichia, Debaryomyces, Yarrowia, Candida* and *Schizosaccharomyces or* selected from the group of moulds belonging to the genera *Aspergillus, Trichoderma, Neurospora, Rhizopus and Penicillium*. Another host cell according to the invention is an insect cell or a plant cell, especially when

the wanted polypeptide undergoes pots-translational modifications which call for the use of insect cells or plant cells. A preferred embodiment of the invention entails the expression of a mammalian polypeptide in relatively simple expression systems that are now equipped with a functional chaperone-like protein specific for the wanted polypeptide. The invention provides a method for the expression, folding, transportation and/or secretion of a wide array of wanted polypeptides, varying from enzymes, hormones, cytokines, growth factors, antibodies, antigens for vaccines and diagnostic tests, and any other polypeptide that may be expressed in expression systems for commercial purposes. The invention provides a method for expressing a polypeptide comprising translating said polypeptide and providing a chaperone-like protein according to the invention. In addition, the invention provides a method for expressing a polypeptide comprising using an expression system or a host cell provided by the invention. Furthermore, the invention provides a method for transporting a polypeptide in a host cell or for secreting a polypeptide from a host cell comprising translating said polypeptide in said host cell and further comprising providing a chaperone-like protein according to the invention or using a nucleic acid, vector, expression system or host cell according to the invention, allowing for the formation of complexes between the chaperone-like protein and (an immature form) of said polypeptide. Furthermore, the invention provides a method according to any wherein said chaperone-like protein is being expressed prior to and/or during the expression of said polypeptide, and, in a preferred embodiment, a method according wherein said chaperone-like protein is being co-expressed with said polypeptide. In addition, the invention provides a method for enhancing correct polypeptide folding comprising creating a mixture which at least comprises said

polypeptide and a chaperone-like protein according to the invention. The method according to the invention is applicable to a polypeptide of mammalian origin and/or to a polypeptide which is post-translationally modified. In the experimental part of the description, the invention provides a method wherein the polypeptide is (pro)chymosine. The invention also provides a chaperone-like protein-polypeptide complex which is obtainable by a method provided by the invention. Such complexes can be isolated from the expression system with methods shown in the experimental part of the description. In nature, proteins (or peptides) with binding characteristics often come in pairs, for example binding between chaperone and corresponding protein, but also for example between ligand and receptor, or peptide and MHC/HLA molecules, or between antigen and antibody. Binding in itself can range from very aspecific on the one hand to very specific on the other hand, as for example is demonstrated when one compares weak hydrophobic interactions between peptide sequences on the one hand with the complex interactions as seen with antibody-antigen binding on the other hand. Binding is dependent on the binding site(s). A binding site can be a simple linear peptide sequence, or a complex of two or more linear sequences comprising a conformational binding site. Another characteristic of protein binding is that it can range from having very low to very high affinity, depending on the 'fit' of the binding site, or binding interaction, or binding motif, or, for example with antigen-antibody binding, the epitope-paratope interaction, between the corresponding peptide sequences and conformational shapes of the two molecules. A chaperone-like protein as provided by the invention is generated by processes and methods from nucleic acid recombinant technology or protein engineering, for example processes resembling DNA or RNA shuffling or random rearrangements

(including mutations, deletions and insertions) or recombinations of DNA or RNA, but can also entail rearrangements or recombinations or mutations, insertions or deletions with or in (parts of) a peptide sequence (constituting a binding motif or paratope or the like) derived from known binding proteins, and such processes each result in a large, in essence unlimited, range or repertoire of (potential) chaperone-like proteins, from which repertoire each resulting chaperone may be specifically well-suited to function in yet another specific heterologous expression system as yet another chaperone for yet another specific protein. Design, construction and selection, can all occur by *in vitro* or *in vivo* methods, or by a combination of these methods. A chaperone-like protein is for instance generated in host cells such as yeast cells (but libraries of cells of other expression systems can also be used) by the generation of host cell banks or libraries in which a nucleic acid sequence (encoding a binding protein or potential chaperone-like protein) is rearranged or shuffled by techniques for example comprising *in vitro* homologous recombination of pools of selected genes (mutant genes and non-mutant genes) by random fragmentation and PCR re-assembly (Stemmer, PNAS 91, 10747-10751), followed by transformation of the host cells with reassembled fragments. Another method to generate a chaperone-like protein is by expressing the peptides or proteins encoded by rearranged nucleic acid sequences (possibly derived from a nucleic acid sequence encoding a binding protein) in a phage display system. This has the added advantage that, depending on the wanted affinity of the chaperone-like protein, phage populations can be selected that are enriched for phages expressing the wanted chaperone-like protein, based on the fact that some phages in the library will react with higher affinity with the corresponding protein than other phages. Yet another possible

method to generate a library of host cells comprising a chaperone-like protein specific for a corresponding protein uses hosts cells comprising nucleic acid molecules of a naturally occurring chaperone, or comprising nucleic acid molecules of an earlier designed or selected chaperone-like protein. Such host cells are modified or transformed by e.g. homologous recombination techniques as above, or directly *in vivo*, with reshuffled or rearranged or in any other way modified nucleic acid, replacing or modifying the known nucleic acid molecules. The invention also provides a method to obtain a chaperone-like protein and host cells comprising a chaperone-like protein by selection for the chaperone-like protein induced rescue of host cells expressing the polypeptide of interest that is inhibitory to growth of these cells, by generating a pool of rearranged nucleic acid molecules encoding binding proteins potentially reactive with said polypeptide, followed by introduction of said pool into host cells capable of producing said polypeptide followed by selecting said transformed host cells for improved growth rates. The invention provides a method for obtaining a transformed host cell comprising a chaperone-like protein comprising transforming a host cell with a nucleic acid encoding a polypeptide to be expressed, or selecting a host cell expressing said polypeptide, further comprising transforming a host cell expressing said polypeptide with a nucleic acid encoding a chaperone-like protein. Such a method provided by the invention to obtain a transformed host cell comprising a functional chaperone-like protein comprises selecting a polypeptide to be expressed in said host cell and transforming a host cell with nucleic acid encoding said polypeptide, or, selecting a host cell capable of expressing said polypeptide, transforming a bank of host cells capable of expressing said polypeptide with a variant nucleic acid molecule derived from a pool of nucleic acid molecules

encoding a potential chaperone-like protein, selecting the thus transformed host cells for improved growth rates and/or increased expression, folding or secretion of said polypeptide. The invention also provides the use of a host cell comprising a chaperone-like protein in the production of a polypeptide, and also provides a chaperone-like protein obtainable from a host cell. These can also be used in a method for improving or enhancing correct *in vitro* or *in vivo* folding of a polypeptide, by creating a mixture which at least comprises a chaperone-like protein provided by the invention and (an immature form of) a polypeptide, or in a method for increasing secretion of a polypeptide from a host cell. The invention is further illustrated by examples in the experimental part of the description without being limited thereto.

Experimental part

EXAMPLE 1 Prochymosin secretion by yeast

As a model system for heterologous protein production we chose prochymosin expression in the yeast *S. cerevisiae*. Mature chymosin is an aspartyl protease that is responsible for the coagulation of milk proteins in the fourth stomach of unweaned calves. It has been a prime target for production in micro-organisms because of its use in cheese production. The mRNA encodes the precursor polypeptide preprochymosin, which is converted into the inactive zymogen prochymosin by removal of the signal peptide in the ER. Native prochymosin is autocatalytically activated to chymosin at low pH. We have expressed prochymosin in yeast (Harmsen, 1995, PhD thesis, Vrije Universiteit, Amsterdam; Harmsen et al., 1996, Appl. Microbiol. Biotechnol. 46:365-370) as a fusion protein consisting of the signal peptide derived from the yeast invertase protein and prochymosin using a suitable plasmid

(pSY78). This plasmid is an integrative *E. coli* -yeast shuttle vector containing a TRP1 gene as the selection marker and an expression cassette consisting of the fusion protein flanked by the GAL7 promoter and the terminator from the PGK1 gene. This plasmid was integrated at the TRP1 locus of strain W303-1A in single copy after linearization with EcoRV, which cuts in the marker gene, resulting in strain MR16. Prochymosin is inefficiently secreted (<1%) by strain MR16. Upon entrance into the acidic culture medium it is activated to chymosin. The intracellularly accumulated prochymosin is not activatable, suggesting that it is misfolded.

EXAMPLE 2 Induction of humoral immune response in a llama

Heavy-chain antibodies are immunoglobulins naturally lacking a light chain; they have been identified in species as varying as camelids and sharks. The variable domains of heavy-chain antibodies (hc-Fv) can be produced in yeast. In order to obtain fragments of heavy-chain antibodies directed against prochymosin a male llama (*Lama glama*) was immunised with authentic mature chymosin. Immunisations were performed both subcutaneously and intramuscularly using 1 ml 0.125 mg/ml rennin (Sigma) per immunisation site. The first two immunisations were performed with a three week interval and using a water in oil emulsion (9:11 (v/v) antigen in water: specol) as described by Bokhout et al. (1981, Vet. Immunol. Immunopath. 2: 491-500). The third immunisation was done without adjuvant five weeks after the first immunisation.

EXAMPLE 3 Construction of libraries for expression of llama hc-Fv fragments

A blood sample of about 200 ml was taken from the immunised llama both at 42 and 77 days after the first immunisation and an enriched lymphocyte population was obtained via Ficoll (Pharmacia) discontinuous gradient

centrifugation. From these cells, total RNA was isolated by acid guanidium thiocyanate extraction, after which first strand cDNA synthesis was performed using random 6-mer primers (Amersham first strand cDNA synthesis kit). DNA fragments encoding hc-Fv fragments and part of either the long or short hinge region were amplified by PCR using respectively primers V_H-2B and BOLI 25 or V_H-2B and BOLI 26:

V_H-2B: 5'-AGGTSMARCTGCAGSAGTCWGG

BOLI 25: 5'-GGAGCTGGGGTCTTCGCTGTGGTGCG

BOLI 26: 5'-TGGTTGTGGTTTTGGTGTCTTGGGTT

The resulting PCR reactions contained predominantly the PCR fragments encoding the hc-Fv. These fragments were digested with *Pst*I and *Bst*EII and ligated into the multiple cloning site of a suitable plasmid, such as plasmid pUR4585 which is a 2μm-derived *E. coli*-yeast shuttle vector suitable for GAL7-driven expression of antibody fragments fused to the invertase signal peptide and containing C-terminal myc and hexahistidine tags. [Plasmid pUR4585 was derived from pSY1 (Harmsen et al., 1993, Gene 125: 115-123) in several steps. First, the *Pst*I site in front of the Gal7 promoter was removed after partial digestion with *Pst*I, incubation with Klenow fragment and subsequent blunt end ligation. Secondly, the *Bst*EII site in the leu2 selection marker was silently removed by site-directed mutagenesis using primers BOLI 3 (5'-CGTTTTGCCAGGCGACCACGTTGGTC) and BOLI 4 (5'-GACCAACGTGGTCGCCTGGCAAAACG). Finally, the about 1.8 kb *Sac*I - *Hin*dIII fragment was replaced with a synthetically prepared fragment (see Figure 1).]

After transformation of *E. coli* strain XL1-blue by electroporation using the four ligations we obtained four libraries, each representing at least 2.5 x 10E5 different colonies. These libraries have been numbered as follows: 42-SH, 42-LH, 77-SH and 77-LH, the number indicating the day at which the blood sample was taken and the text representing

the primer used in the PCR (BOLI 25 is specific for Short Hinge hcFv and BOLI 26 for Long Hinge hc-Fv).

EXAMPLE 4 Isolation of llama hc-Fv fragments specific for prochymosin

The four libraries were introduced into a suitable host cell strain, such as yeast strain W303-1A by transformation according to (Klebe et al., 1983, Gene 25:333-341). At least 2000 colonies of each library were screened for their ability to produce an antibody specific for chymosin using a two membrane system, which is a modification of the procedure described by Skerra et al. (1991, Anal. Biochem. 196:151-155). Briefly, the colonies were transferred from a standard minimal medium plate [0.67% (w/v) Yeast Nitrogen Base without amino acids, 2% (w/v) glucose, 2% (w/v) agar, the relevant amino acids (eacht at 20 mg/l)] to a minimal medium plate containing 2% (w/v) galactose instead of glucose in order to induce expression of GAL7-controlled genes. This plate contained a nitro-cellulose membrane (Optitran BA-S 83 membrane, Schleicher and Schuell) and a polycarbonate membrane. After incubation at 30°C for 72 hr the antibody fragments that were secreted by the yeast cells will have diffused through the non binding polycarbonate membrane and will be captured on the nitro-cellulose membrane. These antibody fragments are then detected by a Western blotting procedure, using biotinylated chymosin (1 μg/ml in PBS containing 0.05% (w/v) Tween-20 and 0.1% (w/v) non fat dry milk) and peroxidase-conjugated streptavidin (Jackson Immunoresearch). Peroxidase staining was performed using Enhanced Chemiluminescence (Amersham). By this method we isolated twelve antibody fragments specific for chymosin, numbers C1 to C96.

The plasmids encoding for these antibody fragments were isolated by plasmid rescue (Robzyk and Kassir, 1992, Nucl. Acids Res. 20: 3790) and after determination of the DNA

sequence of the region encoding the antibody fragments the deduced amino acid sequence of the various antibody fragments was compared. The antibody fragments were clearly derived from clonally unrelated B-cells, as was most evident from the variation in both length and sequence of the CDR3 region.

EXAMPLE 5 Direct selection of host cells comprising a functional chaperone-like protein by suppression of chymosin expression induced growth inhibition

It has been shown previously that overexpression of a heterologous protein that accumulates intracellularly in yeast can be deleterious to growth of the host organism (Romanos et al., 1992, Yeast 8:973-988). Reduction in growth can according to the invention be used to select for host cells such as yeast cells which have an increased secretion efficiency of the heterologous protein and are thus not or less reduced in growth. Intracellular accumulation of prochymosin also results in a decreased growth rate as can be seen when strain MR16 is transformed with the standard shuttle vector YEplac181 (Table 1). When transformants are selected on glucose-plates, which suppresses expression of prochymosin, approximately 4000 transformants form a clearly visible colony, whereas selection of transformants on galactose-plates, with expression of prochymosin, results in a drastically reduced number of transformants that form a colony. However, the host cells which express a chaperone-like protein suppress growth inhibition caused by the prochymosin expression and accumulation, thus growth of these host cells expressing the wanted functional chaperone-like protein on galactose plates must result in the immediate detection of colonies of the wanted transformed host cells. Indeed, when the four libraries are similarly introduced into yeast strain MR16, about 80 transformants obtained with library 77-LH could also form colonies on galactose plates

(Table 1). Six of these transformants were selected for further analysis, numbers C105 to C110.

The plasmids encoding for these antibody fragments were isolated by plasmid rescue and introduced into wild-type yeast strain W303-1A. The binding of the antibody fragments secreted by these transformants to chymosin was then tested by plate-assay. As can be seen in Table 2, clones C108 and C110 were chymosin binders.

EXAMPLE 6 Effect of co-expression of chaperone-like proteins on prochymosin secretion

MR16 producing the various chymosin binding antibody fragments was then tested for the level of chymosin secretion as compared to YEplac181 transformed MR16 using an enzyme assay. As can be seen from Table 2, C96 and C108 clearly result in an increase in chymosin secretion. These clones are also binders for chymosin, suggesting that the binding of the antibody fragments to chymosin was responsible for the increase in the level of chymosin secretion. Many of the antibody fragments that were initially selected for binding to chymosin do not show a chaperone-like effect on prochymosin secretion, possibly because these antibody fragments are specific for epitope(s) that when bound by antibody do not facilitate protein folding. Alternatively, binding by these antibody fragments could possibly inhibit chymosin enzyme activity, masking a stimulatory effect on prochymosin secretion.

EXAMPLE 7 Effect of decreasing the chaperone-like protein expression level on its stimulatory effect on foreign protein secretion

We have previously shown that the stimulatory effect of BiP overexpression on prochymosin secretion by yeast is

increased from 1.5 fold to 26-fold when the level of chaperone overexpression was decreased by using a single copy genomically integrated vector instead of a high-copy number 2μm-based vector encoding the chaperone expression cassette (Harmsen et al., 1996, Appl. Microbiol. Biotechnol. 46: 365-370). Therefore, we similarly deleted the 2μm-based replication origin of plasmids pC96L and pC108L by ligating the 4.5 kb EcoRI-EcoRV fragment of these plasmids to the 0.9 kb EcoRI-EcoRV fragment of YIplac128 (Gietz and Sugino, 1988, Gene 74: 527-534), resulting in pC96R and pC108R, respectively. Transformants of strain MR16 containing these plasmids can be obtained by cutting these plasmids with EcoRV prior to transformation in order to effect integration into the leu2 locus, using the methods described. These transformants, when analysed for their amount of prochymosin secreted into the culture medium as described, clearly show increased chymosin secretion.

EXAMPLE 8 Effect of incorporation of the chaperone-like protein in the yeast cell wall on foreign protein production

Binding of antibody fragments to (pro)chymosin may inhibit the autocatalytic activation of prochymosin and chymosin enzyme activity. In these cases the antibody fragments should be removed from the foreign protein in order to obtain a commercially interesting process. This can for example be accomplished by the display of the antibody fragments on the surface of the cells used for expression (reviewed by Georgiou et al., 1997, Nature Biotechnology 15: 29-34). The foreign protein can then be separated from the antibody fragment by separation of cells from culture medium under conditions which prevent the antibody from binding to its target. The chaperone-like proteins such as antibody fragments C96 and C108 can be incorporated into the host cell wall by fusion, in the case of yeast cells, to the C-terminal fragment of yeast a-agglutinin, a protein normally present in

the yeast cell wall, as has previously been described (Schreuders et al., 1993, Yeast 9: 399-409). This is done by ligating the approximately 350 bp *Pst*I -*Bst*EII fragments of plasmids pC96L and pC108L to the approximately 7.5 kb *Pst*I-*Bst*EII fragment of plasmid pRL25, resulting in plasmids pC96T and pC108T, respectively.

Transformants of MR16 containing either of these plasmids contain the prochymosin bound to the yeast cell wall when they are induced for heterologous protein expression at neutral pH. The chymosin can then easily be purified by collection of yeast cells by centrifugation, washing with PBS, resuspension of cells in 100 mM HCl (pH=1), resulting in dissociation of prochymosin from antibodies and autocatalytic activation of prochymosin to chymosin, and neutralization to pH=7 by addition of 2 M Tris-base.

EXAMPLE 9 Isolation of chaperone-like proteins by chain shuffling

Many chaperones of a particular type are redundant, even within one subcellular compartment. For example, the yeast cytosol contains at least 4 different Hsp70-type chaperones. The yeast ER also contains at least 1 other Hsp70-type chaperone, in addition to the well-known BiP (Baxter et al., 1996, Mal. Cell. Biol. 16: 6444-6456). Not only chaperones but also other types of foldases are redundant. It is assumed that this redundancy functions to cover the range of substrate specificities needed to stimulate the folding of the diverse set of proteins that fold in each compartment. In cases that the specificity recognizing the particular heterologous protein under study is not present in the heterologous host this specificity could be generated by DNA shuffling of the various chaperones of the same class. DNA shuffling (Stemmer, 1994, Nature 370: 389-391) is a method for in vitro homologous recombination of pools of selected

mutant genes by random fragmentation and re-assembly of nucleic acid fragments.

EXAMPLE 10 Removal of chaperone-like protein from secreted foreign protein by introduction of an ER-retention signal

ER-resident proteins contain specific retention signals that distinguish these proteins from proteins that are destined for further transport towards the cell surface. Lumenal (soluble) proteins of yeast usually contain the lumenally-exposed C-terminal sequence KDEL (a related sequence is present in other species, e.g. HDEL in mammals) that is necessary and sufficient for ER-localization (Munro and Pelham, 1987, Cell 48: 899-907). Fusion of the C-terminal sequence KDEL with the antibody fragments C96 and C108 can be accomplished by plasmid construction, such as inserting the appropriate double strand oligonucleotide into the *Bst*EII and *Hin*dIII restriction sites of plasmids pC96L and pC108L. The resulting retention in the ER has the advantage that the concentration of the chaperone-like protein in the ER is higher, which gives a higher potential for chaperone-like protein function. Furthermore, the antibody fragments are not secreted into the culture supernatant together with the foreign protein, where it could inhibit the function of the foreign protein by binding to it.

EXAMPLE 11 Binding site mapping of chaperone-like proteins

In order to determine whether the particular site recognised by the binding site of the chaperone-like protein determines whether it can function as a chaperone it can be determined which of the various specific fragments recognises a different site (epitope) by performing competitive inhibition ELISA (e.g. Kaveri, 1995, Epitope and idiotope mapping using monoclonal antibodies. In: Paul (Ed.), Antibody

Engineering Protocols, Humana Press Inc., Totowa, NJ, PP. 171-181).

EXAMPLE 12 Isolation of chaperone-like proteins from random antibody libraries

A chaperone-like protein could function by binding to a folding intermediate instead of to the authentic correctly folded protein. Since most antibodies obtained after immunisation recognise only the correctly folded protein but not the misfolded protein or folding intermediates, the isolation of chaperone-like proteins by selection of antibodies from immunised organisms may not be optimal. However, in order to obtain a specifically binding antibody fragment it is not necessary to immunise an organism. Large libraries of antibody fragments have been constructed from unimmunised organisms which contain a very diverse set of binding fragments. Antibody fragments binding to any desired target ligand can be isolated from such libraries (Vaughan et al., 1996, Nature Biotechnology 14, 309-314). Such so-called random libraries are constructed by either randomisation of the regions of the variable antibody domains that are involved in antigen binding, the complementarity- determining regions, or by the cloning of a very large number of unselected antibody fragments, or by a combination of these methods. Antibody fragments derived from such random libraries can also have a chaperone-like function.

EXAMPLE 13 Construction of artificial chaperone-like proteins by combinatorial approaches

Novel binding proteins can be obtained by the randomization of surface residues of a parenteral protein molecule that itself does not have the required binding specificity, but which has a known 3D-structure that can be used as a scaffold (reviewed in Nygren and Uhlén, 1997,

Current Opinion in Structural Biology 7:463-469). Such binding proteins can be selected from libraries via specific binding towards a desired target ligand and have the potential to replace natural antibodies or antibody fragments. Such binding proteins have been isolated after randomization of, for example, the Z-domain of protein A (Nord et al., 1997, Nature Biotechnology 15: 772-777) and Tendamistat (McConnell and Hoess, 1995, J. Mol. Biol. 250: 460-470). These artificial binding domains can also be used as a chaperone-like protein. For these purposes the artificial binding domains can be isolated by randomization and expression using a phage display vector in *Escherichia coli* as has been described in the above mentioned references, selection for binding to the desired heterologous protein, followed by selection for specifically binding fragments that have a chaperone-like function. Alternatively, the specifically binding fragment can be directly selected for its chaperone-like function by suppression of heterologous protein expression induced growth inhibition, as was described above.

For the expression of a random library of artificial binding domains of the Z-domain of protein A in yeast a suitable vector, such as plasmid pRL43 can be derived from plasmid pUR4585 in several steps (see Figure 2). First a *Bgl*II restriction site can be introduced into the region encoding the invertase signal peptide by inserting a synthetically prepared double stranded DNA fragment into the *Sac*I and *Pst*I sites of plasmid pUR4585. Secondly, into the *Bgl*II and *Hin*dIII restriction sites of the resulting vector can be inserted a synthetically prepared DNA fragment encoding the Z-domain of protein A and containing suitable restriction sites, which results in the destruction of this *Hin*dIII site and the introduction of a new *Hin*dIII site. Finally, the residues of the Z-domain of protein A whose codons are underlined in Figure 2 can be randomised by performing an assembly PCR using oligonucleotides BOLI A and

BOLI B, digestion of the resulting PCR product with *Sal*I and *Esp*3I and ligating this product to plasmid pRL43 cut with *Sal*I and *Esp*3I.

BOLI A:

5'-TGCGGTCGACAATAAATTCAACAAGGAANNKNNKNNKGCTNNKNNKGAAATTNNKNNKTTGCCCAACTTGAAC

BOLI B:

5'-TGGCCGATTGAGACGGATCATCMNNCAAGCTMNNGATAAAAGCMNNMNNTTGMNNMNNGTTCAAGTTGGGCAA

After amplification of the resulting random library in *E. coli* it can be used to transform yeast cells expressing a heterologous protein that is deleterious for cell growth, such as strain MR16 (see above) using an efficient transformation protocol, such as lithium acetate-transformation (Ito et al., 1983, J. Bacteriol. 153: 163-168). Yeast cells containing the wanted chaperone-like protein can then be selected as described above.

Legends to the figures

Figure 1. Insert of plasmid pUR4585 which is suitable for expression of a fusion protein between the yeast invertase signal peptide (single underlined), a llama heavy-chain antibody fragment (italicized), a myc tail (bold) and a hexahistidine tail (double underlined) in yeast. Heavy-chain antibodies should be inserted as *Pst*I and *BstE*II restriction fragments into this vector. These restriction enzymes cut in the beginning and end of heavy chain antibodies. Therefore, the first 5 amino acids (sequence QVQLQ) and the last 6 amino acids (sequence QVTVSS) are encoded by this vector. Only relevant restriction sites are indicated.

Figure 2. Insert of plasmid pRL43 encoding a fusion protein consisting of the invertase signal peptide (underlined) and the Z-domain of protein A. The codons for the amino acids that are used for randomisation are underlined. The former HindIII restriction site of plasmid pUR4585 that was destroyed in the construction of pRL43 is indicated in bold.

Table 1. Suppression of growth defect of chymosin producing yeast strain MR16 after transformation with libraries encoding chymosin specific hc-Fv.

	number of MR16 transformants forming a colony[a]	
library	glucose	galactose
YEplac181	4000	2
42-SH	2000	2
42-LH	1000	1
77-SH	2000	1
77-LH	2000	80

[a] After transformation of yeast strain MR16 with the indicated libraries or control plasmid YEplac181 the cells were spread on plates containing either glucose, which suppresses chymosin expression, or galactose, which induces chymosin expression.

Table 2. Chymosin binding and effect on chymosin secretion of various antibody fragments.

C-clone	library	chymosin binding of W303-1A transformants[a]	relative chymosin activity of MR16 transformants[a]
1	77-LH	+	+ -
5	77-LH	+	-
6	77-LH	+	-
9	77-LH	++	-
13	77-LH	+	-
14	77-LH	+	-
15	77-LH	+	-
31	77-LH	+	-
34	77-LH	+	+ -
35	77-LH	++	+ -
95	42-LH	+	-
96	42-LH	++	+
108	77-LH	+	+
110	77-LH	+	+ -

[a] The signal strength relative to yeast cells transformed with control plasmid YEplac181, which does not encode an antibody fragment, is indicated. ++, strongly increased signal; +, increased signal; + -, possibly increased signal; -, no increase.

CLAIMS

1. A chaperone-like protein having been provided with at least one binding site with binding activity for at least one immature form of a polypeptide for which it originally had no, less or different binding activity.

2. A chaperone-like protein with a binding site according to claim 1 which binding site corresponds to a peptide sequence binding with at least one immature form of said polypeptide.

3. A chaperone-like protein with a binding site according to claim 1 or 2 which binding site is derived from a binding protein selected from the group of enzymes, receptor molecules, cell surface proteins, chaperones and transfer proteins.

4. A chaperone-like protein with a binding site according to claim 1 or 2 which binding site is derived from a binding protein selected from the group of antibodies, antibody fragments, single-chain antibody fragments and variable regions of the heavy or the light chain of antibodies.

5. A chaperone-like protein according to claim 4 wherein the binding protein is selected from the group of camelid heavy-chain antibodies or fragments thereof.

6. A chaperone-like protein according to claim 1, 2, 3, 4 or 5 wherein said chaperone-like protein specifically binds with (pro)chymosin.

7. A nucleic acid encoding a chaperone-like protein according to any of claim 1 to 6.

8. A nucleic acid according to claim 7 additionally comprising a nucleic acid sequence encoding said polypeptide.

9. A vector comprising a nucleic acid according to claim 7 or 8.

10. A vector according to claim 9 additionally comprising nucleic acid sequences allowing for controlled expression of the chaperone-like protein and/or said polypeptide.

11. An expression system comprising a nucleic acid molecule according to claim 7 or 8 or a vector according to claim 9 or 10.

12. A host cell comprising a nucleic acid molecule according to claim 7 or 8 or a vector according to claim 9 or 10.

13. A host cell according to claim 12 which is selected from the group of bacteria belonging to the genera *Escherichia, Bacillus, Streptomyces and Lactobacillus.*

14. A host cell according to claim 12 which is selected from the group of yeasts belonging to the genera *Saccharomyces, Kluyveromyces, Hansenula, Pichia, Debaryomyces, Candida, Yarrowia* and *Schizosaccharomyces* or from the group of moulds belonging to the genera *Aspergillus, Trichoderma, Neurospora, Rhizopus and Penicillium.*

15. A method for expressing a polypeptide comprising translating said polypeptide and providing a chaperone-like protein according to any of claims 1 to 6.

16. A method for expressing a polypeptide comprising using an expression system or a host cell according to any of claims 11 to 14.

17. A method for transporting a polypeptide in a host cell or for secreting a polypeptide from a host cell comprising translating said polypeptide in said host cell and further comprising providing a chaperone-like protein according to any of claims 1 to 6 or providing a nucleic acid according to claim 7 or 8, or providing a vector according to claim 9 or 10, allowing for the formation of complexes between the chaperone-like protein and (an immature form) of said polypeptide.

18. A method according to any of claim 15 to 17 wherein said chaperone-like protein is being expressed prior to and/or during the expression of said polypeptide.

19. A method according to any of claims 15 to 17 wherein said chaperone-like protein is being co-expressed with said polypeptide.

20. A method for enhancing correct polypeptide folding comprising creating a mixture which at least comprises said polypeptide and a chaperone-like protein according to any of claims 1 to 6.

21. A method according to any of claims 15 to 20 wherein the polypeptide is of mammalian origin.

22. A method according to any of claims 15 to 21 wherein the polypeptide is post-translationally modified.

23. A method according to claim 20 wherein the polypeptide is (pro)chymosin.

24. A chaperone-like protein-polypeptide complex obtainable by a method according to any of claims 15 to 23.

25. A method for obtaining a transformed host cell comprising a chaperone-like protein comprising transforming a host cell with a nucleic acid encoding a polypeptide to be expressed, or selecting a host cell expressing said polypeptide, further comprising transforming a host cell expressing said polypeptide with a nucleic acid encoding a chaperone-like protein.

26. A host cell obtainable by a method according to claim 25.

27. A method for expressing a polypeptide comprising using a host cell according to claim 25.

28. Use of a method according to any of claims 15 to 22 or 26 or a host cell according to any of claims 11 to 15 or 27 in the production of a polypeptide.

29. A polypeptide obtainable by a method according to any of claims 15 to 23 or 27.

30. A polypeptide obtainable from a chaperone-like protein-polypeptide complex according to claim 24.

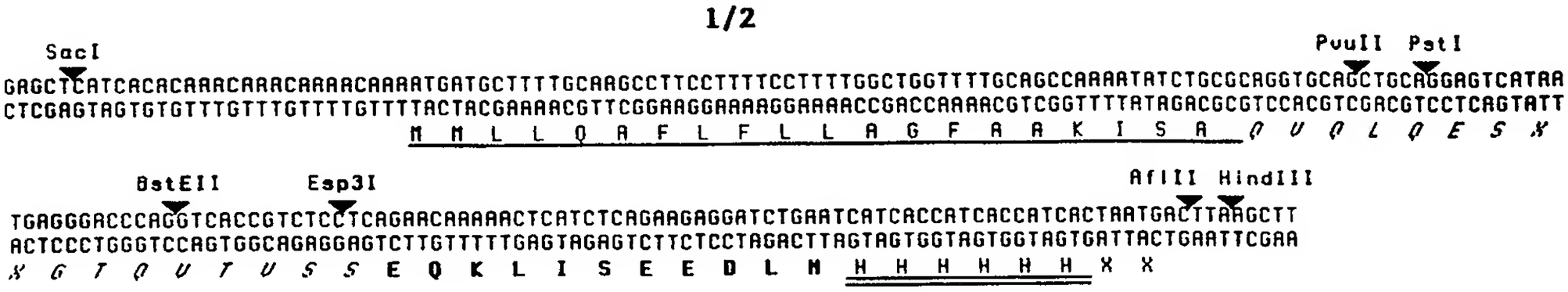

Figure 1. Insert of plasmid pUR4585 which is suitable for expression of a fusion protein between the yeast invertase signal peptide (single underlined), a llama heavy-chain antibody fragment (italicized), a myc tail (bold) and a hexahistidine tail (double underlined) in yeast. Heavy-chain antibodies should be inserted as *Pst*I and *Bst*EII restriction fragments into this vector. These restriction enzymes cut in the beginning and end of heavy chain antibodies. Therefore, the first 5 amino acids (sequence QVQLQ) and the last 6 amino acids (sequence QVTVSS) are encoded by this vector. Only relevant restriction sites are indicated.

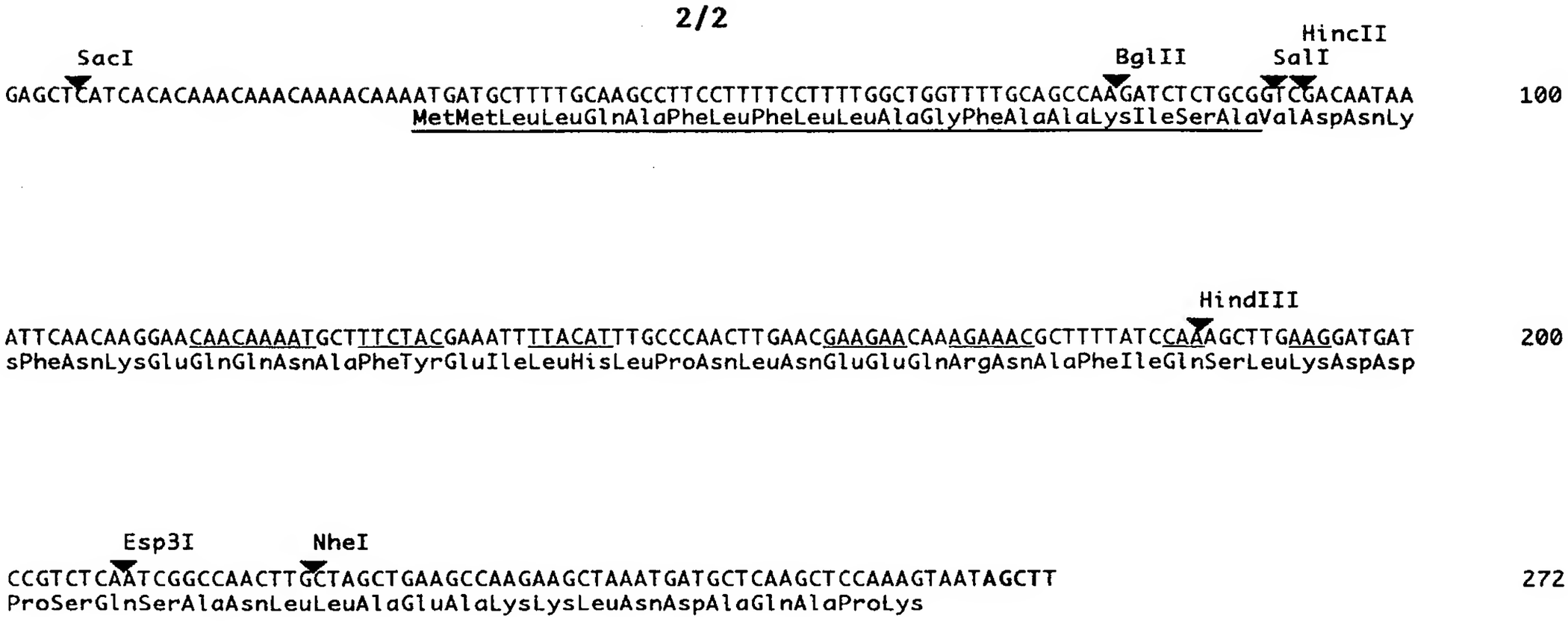

Figure 2. Insert of plasmid pRL43 encoding a fusion protein consisting of the invertase signal peptide (underlined) and the Z-domain of protein A. The codons for the amino acids that are used for randomisation are underlined. The former HindIII restriction site of plasmid pUR4585 that was destroyed in the construction of pRL43 is indicated in bold.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/NL 98/00335

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/62 C07K19/00 C12P21/02 C12N15/70 C12N15/74
C12N15/81 C12N15/80 C12N1/21 C12N1/15 C12N1/16

According to International Patent Classification(IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K C12P

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	GETHING M -J ET AL: "BINDING SITES FOR HSP70 MOLECULAR CHAPERONES IN NATURAL PROTEINS" COLD SPRING HARBOR SYMPOSIA ON QUANTITATIVE BIOLOGY, vol. 60, 1995, pages 417-428, XP000613691 see the whole document ---	1-30
A	HARMSEN, M. M. ET AL: "Overexpression of binding protein and disruption of the PMR1 gene synergistically stimulate secretion of bovine prochymosin but not plant thaumatin in yeast" APPL. MICROBIOL. BIOTECHNOL. (1996), 46(4), 365-370 CODEN: AMBIDG;ISSN: 0175-7598,1996, XP002045333 see the whole document -----	1-30

☐ Further documents are listed in the continuation of box C.

☐ Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
27 October 1998	02/11/1998

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Montero Lopez, B

1

Form PCT/ISA/210 (second sheet) (July 1992)